



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/545,078	04/07/2000	Stephane H Maes	Y0999-273(8728-298)	1114
46069 7590 12/19/2006 F. CHAU & ASSOCIATES, LLC 130 WOODBURY ROAD WOODBURY, NY 11797			EXAMINER PRIETO, BEATRIZ	
			ART UNIT 2142	PAPER NUMBER

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
2 MONTHS	12/19/2006	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.



UNITED STATES PATENT AND TRADEMARK OFFICE

---

Commissioner for Patents  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

**BEFORE THE BOARD OF PATENT APPEALS  
AND INTERFERENCES**

Application Number: 09/545,078  
Filing Date: April 07, 2000  
Appellant(s): MAES, STEPHANE H

**MAILED**

**DEC 19 2006**

**Technology Center 2100**

---

Frank V. DeRosa (Reg. No. 43,584)  
For Appellant

**EXAMINER'S ANSWER**

This is in response to the appeal brief filed 09/19/06 appealing from the Office action mailed 12/09/06.

Art Unit: 2142

**(1) Real Party in Interest**

A statement identifying by name the real party in interest is contained in the brief.

**(2) Related Appeals and Interferences**

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal except from those set forth by appellant.

**(3) Status of Claims**

Although Appellant's statement of the status of all the claims in the proceeding (e.g., rejected, allowed or confirmed, withdrawn, objected to, canceled) is correct, however, the statement seems to inadvertently provide a statement and/or identification of those claims that are being appealed.

Since, the record is clear with this respect and the presented Appendix of the claims involved in the appeal is correct, claims 1, 3-15, 17-19, 21-23, 25-34 and 36-37 are those claims being appealed.

**(4) Status of Amendments After Final**

The statement of the status of claims contained in the brief is correct.

**(5) Summary of Claimed Subject Matter**

The summary of claimed subject matter contained in the brief is correct.

**(6) Grounds of Rejection to be Reviewed on Appeal**

The appellant's statement of the grounds of rejection to be reviewed on appeal is hereby corrected in view of this Examiner's answer. Namely, due to the withdrawal of rejection of claims 1, 23 and 34

Art Unit: 2142

under 35 USC 112, second paragraph enclosed herein, Appellant's statement (on page 6 of brief) regarding VI. Grounds of rejection to be reviewed on appeal should no longer contain reference to this rejection.

**(7) Claims Appendix**

The copy of the appealed claims contained in the Appendix to the brief is correct.

**(8) Evidence Relied Upon**

The following evidence/references are relied upon by the examiner in the rejection of the claims under appeal.

- |    |              |                 |               |
|----|--------------|-----------------|---------------|
| 1. | US 6,859,776 | Cohen et. al.   | Feb. 22, 2005 |
| 2. | US 6,501,832 | Saylor et. al.  | Dec. 31, 2002 |
| 3. | US 5,614,940 | Cobbley et. al. | Mar. 25, 1997 |

**(9) Grounds of Rejection**

1. Quotation of the appropriate paragraphs of 35 U.S.C. 103 that form the basis for the rejections under this section made in this Office action may be found in previous office action.
2. Claims 1, 3-15, 17-19, 21-23, 25-34, 36-37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Saylor et. al. (US 6,501,832) (Saylor hereafter) in view of Cohen et. al. (US 6,859,776) (Cohen hereafter) in further view of Cobbley et. al. (US 5,614,940)

Regarding claim 1, Saylor teaches a system comprising

a ("conversational") browser, which provides a ("conversational") user interface to enable access to a ("conversational") portal (12) across a ("plurality of different modalities") two or more multi-

Art Unit: 2142

modal dialog (Saylor: speech-only or text via a phone or Web, respectively see col 4/lines 43-col 5/line 20);

wherein the conversational browser establishes or supports ("adapts") an interaction dialog between the conversational portal (12) and a client (14) based on one modality supported by the client e.g. audio or visual capability (access over the phone or logging via Web see col 4/lines 43-col 5/line 8, retrieve voice or visual display of requested content see col 5/lines 9-20, text for visual display see col 8/lines 32-36);

wherein the conversational browser can retrieve one information "page" from an information source (70, 18 or 22) in response to a client's request and server or present the retrieved page to the requesting client in a format that is compatible with one modality of the client by converting the retrieved page, if necessary to a format compatible with the client (col 8/lines 14-63, col 5/lines 4-8, and col 5/lines 9-20, text-speech conversion for client using a phone see col 4/lines 16-28, and col 8/lines 29-32);

said plurality of modalities including audio (e.g. voice) and non-audio modalities (e.g. text) (col 5/lines 9-20, col 8/lines 14-36, and col 12/lines 20-44), although Saylor does not explicitly teach information ("page") that can be rendered in two modalities.

Cohen teaches a page configure with two modalities, specifically, both graphical and voice data (col 10/lines 39-55) or configured with format supporting both voice and data (col 9/lines 32-43), associated with a browser configured to rendered respective two modalities (col 9/lines 10-25);

It would have been obvious to one ordinary skilled in the art at the time the invention was made given the teachings of Saylor configuring portal with voice browser enabling access to clients across two modalities, the teachings of Cohen for enhancing voice browser with dual capabilities supporting two modalities voice and visual data, would have been readily apparent Motivation would be to configure the browser with gateway capabilities to couple audio voice and conventional Web page information enabling user to voice from voice providers and non-voice data from Internet providers, enabling access to web pages having two modalities, such as those pages configured with markup languages that support two modalities or pages that can be converted from one modality to another for rendering, as suggested by Cohen.

Saylor further teaches providing access of said page to the user on a subscription basis to the content (col 6/lines 48-65, see registration/subscription module 7 of Figs. 2 and 7), upon user request (col 10/lines 47-61, user input request for particular content then request col 11/lines 33-43, also retrievable upon demand, e.g. via an internet protocol col 11/lines 43-52);

Art Unit: 2142

where the user registers/subscribed with the provider to obtain access to the content (col 7/lines 56-col 8/line 12) stored in an user accessible database (18) (col 2/lines 12-17, and col 14/lines 19-30), including audio content, text and (col 4/lines 43-col 5/line 20);

said database comprising indexed content (col 18/lines 17-23, col 8/lines 19-41), access and retrieval of desired content stored on the database is upon user request (col 10/lines 47-61, user input request for particular content then request col 11/lines 33-43, also retrievable upon demand, e.g. via HTTP protocol col 11/lines 43-52);

portal (12) comprises a plurality of modules for providing content to the user for the received user input (col 18/lines 8-12), said modules (i.e. programs) including specifically,

receiving user's input request (col 10/lines 37-46) is provided to a program (VCode identifier module) to identify the Vpage and its location. The identification of the desired page is passed to a program (VPage retrieval module), which accesses a database in communication with the VNAP to retrieve the requested content, the retrieved content is then passed to a program (VPage execution module and VPage menu module). The VPage execution module executes the VPage in conjunction with a voice output module, the voice output module generates audible output from the VPage under control of the VPage execution module and passes the audible output to the call center to relay it over the communications network to the requesting user (col 10/lines 47-61); although Saylor teaches where the databases store audio and non-audio data accessible by the subscriber, including storing the indexed audio in an index database (Saylor: col 8/lines 37-53, index database; col 18/lines 13-19), Saylor does not teach where this data is segmented;

Cobbley teachings pertaining access to information over a network, teaches system (100) for segmenting and indexing audio and non-audio data from a source (105) (col 3/lines 26-col 4/line 7), accessible upon request (col 8/lines 24-29) by a user, specifically broadcast information is provided upon demand, i.e. upon request (col 15/lines 19-32).

It would have been obvious to one of ordinary skill in the art at the time the invention was made given the teachings of Saylor for providing content to the client upon demand, the teachings of Cobbley would have been readily apparent. Specifically, given that content provided by a subscription based service to the user upon his/her request in the Saylor reference is indexed and comprises audio and non-audio data one would be motivated by Saylor's suggestion where the VNAP portal may also maintain an index of information or other content that is available corresponding to a user's request, e.g. storing an index of other information (in other formats), and in addition thereto, text or other content may be displayed on or output to the user's phone or other terminal device (sic), the content generated on the fly

Art Unit: 2142

from other content e.g. news also provided in the Cobbley provides text that is used for indexing non-audio data in doing so non-audio data can be requested by subject matter as input, this text is indexing information which can be transmitted along audio broadcast along with the non-audio broadcast data such as video, as taught by Cobbley.

Claim 2 (canceled)

Regarding claim 3, the information provided by the information sources is implemented in a multi-modal representation, which is a format (Saylor: col 14/lines 46-60, multi-modal content col 23/lines 34-47, col 4/lines 43-col 5/line 20).

Regarding claim 4, computing device (transcoder), operatively associated with the conversational browser, for converting the (multi-modal) information into one specific format (modality-specific format) that corresponds to a modality of the requesting client (Saylor: col 23/lines 60-67, col 21/lines 7-41).

Regarding claim 5, the conversational portal discovers, ascertain, identify (i.e. detects) the modality of the requesting client to convert the multi-modal information into the modality-specific format (Saylor: col 21/lines 7-41, modality of client: col 18/lines 45-col 19/line 11, col 19/lines 46-50, conversion: col 23/lines 60-67).

Regarding claim 6, discern (i.e. detects) the modality of the requesting client based on the registration protocols (Saylor: registration setup: transactions based on user identified registration, col 8/lines 56-col 8/line 13, subscribe registration discern modalities of client, col 16/lines 18-col 17/line 4).

Regarding claim 7, comprising a (portal directory) database (Saylor: 18 of Fig. 1, col 14/lines 19-28), accessible by the conversational browser, for storing one of an index of information sources, (Saylor: indexed: col 18/lines 13-19, index: col 8/lines 14-36).

Regarding claim 8, the information is maintained in a (multi-modal) format by a service provider (Saylor: 70 of Fig. 1, col 11/lines 43-52) of the conversational portal under business agreements between the service provider of the conversational portal and service provider of the information sources (Saylor: business arrangement: col 34/lines 14-34, business: col 38/lines 13-24).

Art Unit: 2142

Regarding claim 9, capturing a connection between the requesting client and the conversational portal and maintaining communication link (holding the client captive) during predetermined time periods (Saylor: col 26/lines 39-40).

Regarding claim 10, a link provided by the conversational browser is selected by the requesting client and rendered or served to the requesting client (Saylor: voice browser; col 26/lines 41-54).

Regarding claim 11, the requesting client is released when a link is directly requested by the requesting client (Saylor: col 26/lines 39-40).

Regarding claim 12, a service provider of the conversational portal provides advertisements, during time period in which the requesting client is connected (Saylor: col 26/lines 39-40).

Regarding claim 13, time period between connections established links between different information sources (Saylor: col 26/lines 39-40).

Regarding claim 14, wherein the advertisements and services are multi-modal (Saylor: content: col 2/lines 5-16, visual or audio content, & multi-modal information col 4/lines 16-28, advertisements: col 9/lines 54-58, ads; col 36/lines 48-67).

Regarding claim 15, wherein the advertisements and services are provided by the service provider on behalf of a third-party under a business agreement between the service provider of the conversational portal and third-party (Saylor: col 36/lines 48-col 37/line, third-party, col 45/lines 65-col 46/line 25).

Claim 16 (canceled)

Regarding claim 17, retrieving as discussed above, obtains desired segments from the multimedia database in response to a client request and presents such segments to the client (Saylor: col 8/lines 14-53, index database; col 18/lines 13-19) based on the I/O capabilities of the client (Saylor: col 3/line 63-col 4/line 4 format that is compatible with the I/O modalities of the requesting client col 8/lines 14-36, presented to the user: col 5/lines 4-8, provide to user visual or audio content: col 5/lines 9-20).



Art Unit: 2142

Regarding claim 18, the conversational browser periodically downloads multimedia data from one information source to index and store the multimedia data in the multimedia database (Saylor: col 18/lines 32-38).

Regarding claim 19, the downloading and indexing of the multimedia data of the information source is performed under a business agreement between a service provider of the conversational portal and a service provider of the information source (Saylor: col 7/lines 4-17).

Claim 20 (canceled)

Regarding claim 21, (registered) subscriber can conversationally navigate the program and select desired segments for broadcasting via the requesting client (Saylor: col 5/lines 9-20).

Regarding claim 22, however the above-mentioned prior art of record does not teach radio services upon request (i.e. upon demand) which the registered subscriber access via a wireless phone (Saylor: wireless access protocol device see col 5/lines 9-20, wireless network providers accessed over wireless phones see col 5/lines 31-37, over a wireless telephone network see col 10/lines 6-13).

Regarding claim 23, comprising features discussed on claim 1, same rationale of rejection is applicable and further:

an access device having one modality (Saylor: device supporting one modality: col 9/lines 59-col 10/line 5, client access device: col 5/lines 9-20 supporting one modality, user device: col 14/lines 10-14);  
a content server (70 of Fig. 1, col 11/lines 43-52 or 18 of Fig. 1, col 14/lines 19-28).

Claim 24 (canceled).

Regarding claim 25, wherein the multi-modal content pages and applications are implemented in a modality-independent representation (Saylor: col 4/lines 43-59, multi modal, col 14/lines 56-59, multi-modal Vpages, col 21/lines 7-41).

Regarding claim 26, a voice (conversational) browser for fetching and processing a multi-modal content page for presentation to a user based on the I/O capabilities of the access device supporting user selection (Saylor: processing for presentation, i.e. rendering see col 26/lines 41-45, fetch col 29/lines 17-28).

Art Unit: 2142

Regarding claim 27, converting one multi-modal content page into one modality-specific representation based on detected modality of the access device (Saylor: col 23/lines 60-67, col 21/lines 7-41).

Regarding claim 28, accessible by the conversational portal, for converting a modality-specific site of a content provider into a multi-modal representation (Saylor: the conversational portal discover, ascertain, identify (i.e. detects) the I/O modalities of the requesting client to convert the multi-modal information into the modality-specific format col 21/lines 7-41, modality of client: col 18/lines 45-col 19/line 11, col 19/lines 46-50, conversion: col 23/lines 60-67).

Regarding claim 29, the conversion service is provided by the content provider (Saylor: col 2/lines 21-26, 59-64) or a third-party under a business agreement with the content provider (Saylor: col 7/lines 4-23).

Regarding claim 30, an interface (conversational browser) (Saylor: interface col 5/lines 2-8) for processing and presenting one of a multi-modal content page and application received by the conversational portal (Saylor: col 10/lines 62-64).

Regarding claim 31, the conversational portal is access by calling a pre-designated telephone number (Saylor: dialing a number, col 6/lines 30-33, connect via telephone, col 25/lines 59-63).

Regarding claim 32, a listing (directory) of content server addresses (Saylor: col 14/lines 23-32).

Regarding claim 33, wherein the (portal directory) database comprises one of multi-modal content pages (Saylor: col 14/lines 23-27), associated with one content provider hosted by the conversational portal (Saylor: multi-modal pages associated with provider 70 of Fig. 1, col 11/lines 43-52)

Regarding claim 34, comprising features discussed on claims 1 and 23, same rationale of rejection is applicable, and further wherein: access to information is over a communication network (16 of Fig. 1) (Saylor: col 14/lines 14-22); establishing communication with a conversational portal using an access device (14 of Fig. 1) (Saylor: col 14/lines 1-22, access via communication link, col 5/lines 21-37, communication medium, col 11/lines 43-52).

Claim 35 (canceled).

Art Unit: 2142

Regarding claim 36, detecting by the conversational portal at least one modality of the access device (Fig. 8), and transcoding i.e. converting the retrieved multi-modal information into at least one modality-specific format corresponding to the at least one detect modality (Saylor: text-speech conversion for client using a phone see col 4/lines 16-28, and col 8/lines 29-32).

Regarding claim 37, established communication with the user (i.e. holding the user captive) during a period in which the retrieving step is executed, presenting one of advertisements (Saylor: col 26/lines 39-40).

Claim 38 (canceled).

**(10) Response to Arguments**

**A.1 Regarding claims 1 and 23, rejected as being unpatentable over Saylor in view of Cohen in further view of Cobbley, it is argued (p. 8-10 of brief) that the applied references do not teach:**

(i) a “conversational browser”, because the Cohen reference which was relied on for teaching “*a conversational browser that can present or serve a multi-modal page having a multi-modal format that can be rendered in two or more synchronized modalities*”, according to Appellant, teaches that hypertext links to content can be presented via audio or by display as disclosed on column 10, lines 39-55. It is further, not clear to Appellant where in the Cohen reference the browser (10) can process a content page in multi-modal format that can be rendered in two or more synchronized modalities. Because according to Appellant (on p. 11 of brief), that the *hypertext lines links in speech or audio to fetch content pages in the reference are “starkly different” than a multi-modal documents rendered by a conversational browser in two or more synchronized modalities*.

In response to the above-mentioned argument, Appellant’s interpretation of the applied reference has been fully considered. Claim terminology has been reviewed to ensure the proper interpretation to the claimed language has been applied. In this case, the specification has been reviewed for a deliberate, explicit definition to the claimed term “*conversational browser*” that will control the interpretation of the term as it is used in the claim. Appellant have pointed to Fig. 4, page 20 to page 22, line 12 (on p. 3 and 4 of brief). The cited portions have been reviewed.

The disclosure of page 20-22, describes “Although any suitable multi-modal browser may be implemented in the conversational portal 11, one preferred architecture for the portal conversational browser 22 is illustrated in Fig. 4 and described in detail in the above-incorporated US serial 09/507,526.

Art Unit: 2142

Briefly, as illustrated in Fig. 4, a *conversational (multi-modal) browser 40* comprises a plurality of mono-mode browsers (e.g. a visual browser 44 (HTML) and a speech browser (45) (VoiceXML) as shown), a multi-modal shell API 41 and a multi-modal shell 42 having a registration table 43 (the multi-modal shell 42 executes on top of any conventional operation system/platform)....". Further, the disclosure of page 22 described Figure 5 as illustrating "*another preferred architecture for the portal conversational browser 22 that may be employed* utilizing a conversational virtual machine (CVM) when more complex conversation computing features are required, etc." Figure 4, as described by "Brief Description of the Drawings" is described as: "Fig. 4 is a block diagram of an *architecture of conversational (multi-modal) browser that may be employed* in connection with the present invention; and Fig. 5 is a block diagram of an architecture of *another conversational (multi-modal) browser that may be employed* in connection with the present invention, see page 10 of the specification.

Appellant has indicated (see p. 10 of brief) that quote "Fig. 4 of the specification illustrates an *exemplary embodiment* of a "conversational" browser comprising a speech browser and visual browser with a browser shell that coordinates the synchronization between modalities (see e.g. page 20-22 and pages 26-28 of the specification)."

Thus, the written disclosure has not set forth a "with reasonable clarity, deliberateness, and precision" a controlling definition to the claimed term "conversational browser" other than "any suitable multi-modal browser may be implemented in the conversational portal 11" and a "*preferred architecture for the portal conversational browser*" described with respect to Figures 4-5.

Thus, because there is no explicit, deliberate and precise definition for the claimed term "conversational browser" in the specification, this term has been given its ordinary meaning and broadest reasonable interpretation (i.e. a browser) and has not been limited to preferred embodiment(s) in the specification as required by MPEP 2111.01, indicated and relied on by Appellant.

Further, the specification has been reviewed for a deliberate, explicit definition to the claimed term "synchronized modality" that will control the interpretation of the term as it is used in the claim. Appellant have pointed to quote "(see e.g. specification page 15, line 2 to page 19, line 7 and page 24, line 13 to page 28, line 18)" on summary of claimed subject matter on p. 3 of brief. Page 15-16 discloses where the term "modality" seems to be interchangeable with "capability of the device processing or rendering the content", page 16, lines 9-10; "multi-modal documents (such as gesture-based conversational markup language (CML) documents) to an access device running a conversational (multi-modal) browser to processing/rendering by the local conversational browser", see page 17, lines 7-11; "the implementation of multi-modal documents, which can be transformed to document of desired

Art Unit: 2142

modalities, ensures content is reused and meets the accessibility requirements (e.g. a multi-modal document designed with combined *visual and aural* modalities can be used in environments where only one modality is available)” (see page 19, lines 1-8 of specification). Synchronized tags are used to synchronize different markup language ML content (i.e. to synchronize an action of a given command in one modality with corresponding actions in the other supported modalities) on an element-by-element basis. The techniques are described, for example, in the above-mentioned incorporated application US serial No. 09/507,526 filed on February 18, 200 (see page 17, lines 15 to page 18, line 8. Page 21, and 26 of the written disclosure as filed mentions modalities that are synchronized, fails to set forth an explicit, deliberate and precise definition for the claimed term “*synchronize modalities*” that will control the interpretation of the term as it is used in the claim. On page 27 of the written description the claimed term “*synchronized modalities*” are exemplified as WML and VoiceXML documents that are synchronized for a multi-modal WAP browser (i.e. micro-browser for the WML modality (see page 27, lines 15-23).

Thus, the written disclosure has not set forth “with reasonable clarity, deliberateness, and precision” a controlling definition to the claimed term “*synchronized modalities*” other than an exemplified embodiments discussed above, pointed out by Appellant.

Thus, because there is no explicit, deliberate and precise definition for the claimed term “*synchronized modalities*” this term will be given the broadest reasonable interpretation (i.e. modalities) and will not be limited to preferred embodiment(s) in the specification (see MPEP 2111.01), as argued.

Hence, argued claim limitation, *a conversational browser that can present or serve a multi-modal page having a multi-modal format that can be rendered in two or more synchronized modalities*”, will be interpreted as,

*a browser presenting a page rendered in two or more modalities.*

Cohen teaches a browser presenting a page rendered in two modalities. Specifically, “A connection to a Web page which includes audio capability, or plain text that can be read via a text-to-speech converter. Such a **page** is configured to provide either graphical data, voice data, or **both**, depending upon the type of equipment that accesses it. In this way, links can be displayed graphically as **hypertext links** in the usual way or as **voice links** with an earcon, or other audio indication, or **both**. Certain links will only be available to a computer user logged onto the Internet and only provide graphical information. Such links will not be presented with an earcon to the originating user. Other links will only be to voice services and will only provide audio information. Such links will not be presented with a hypertext link on the

Art Unit: 2142

graphical Web page. Still **other links** will be to a **data provider** that offers **both graphical and audio data** and **both** an audio indication and a hypertext link will be available (see Cohen column 10, lines 39-55). The Web page is access via the browser controller through a modem via a gateway (see Cohen column 10, lines 57-61).

Appellant's arguments that neither the references teaches a "conversational (multimodal) browser", have been fully reconsidered but not found persuasive. Appellant's arguments that the browser (10) in the Cohen reference does not teach "synchronized" modalities has been fully considered, however, given the broadest reasonable interpretation the applied references does teach, where browser (10) presents a page rendered in two modalities (e.g. graphically displayed hypertext links and voice links or both graphical and audio data and both an audio indication and a hypertext link will be available).

Arguments that in the Cohen reference that hypertext links to content can be presented via audio or by display as disclosed on column 10, lines 39-55 have been considered but that interpretation is inconsistent with what the portion really discloses. The reference does clearly disclose presenting a page with BOTH audio and visual data.

Arguments, that Cohen does not render or present a page rendered in two or more synchronized modalities has been considered, however, given the broadest reasonable interpretation in light of the disclosure as mandated of the term "synchronized modalities", the reference reads on the claim language as interpreted by the cited portions above.

**A.1 Regarding claims 1 and 23, rejected as being unpatentable over Saylor in view of Cohen in further view of Cobbley, it is argued (p. 9 of brief) that the applied references do not teach:**

(i) a "conversational browser", because it is not clear what component of portal (12) in Figures 1 or 2 of the Saylor reference is a conversational browser.

In response to the above-mentioned argument, according to the invention's written specification with respect to the claimed "conversational portal system", a portal may be a gateway see specs page 3, lines 10-12; a conversational portal comprises a conversational browser which retrieves information from an information source in response to a request from a client and then server or presents the retrieved information in a format that is compatible with the I/O modality of the requesting client see specs page 7, lines 17 to page 18, line 6, *this is the function of a proxy server*. Further according to applicant's specification, the conversational portal comprises a portal proxy/capture module (20 on Fig. 1) see specs

Art Unit: 2142

page 19, lines 8-10. System 10 on Fig. 1 further includes a conversational proxy server (27) having a transcoder (28) which may be used to transcode pages/application of one or more sites of a given content provider see specs page 28, line 19 to page 29, line 5.

According to the claim limitation "*a computing system comprising a conversational browser to provide a user interface that adapts an interaction dialog between the conversation portal and a client according to one more modalities*".

Although there seems to be a typographical error with respect to "*one more modalities*", for the purposes of examination, this has been taken as reading "*one or more modalities*".

Saylor discloses a system that enables a user to access content corresponding to the Vcode (i.e. code assigned to a particular page of content delivered to a user via a phone or other delivery device) by entering the VCode through an interactive communications device, such as a land or mobile telephone. The use interacts with the system via inputted verbal descriptions which once interpreted by the system are then associated with the desired VCode and corresponding pages or pages of content to generate for the user (see column 1, line 55 to column 2, line 3).

The VNAP receives the telephone call from the user, *prompts* the user for a VCode corresponding to desired information, and then *delivers that desired* information. The VNAP may comprise a number of modules to accomplish this functionality. According to one embodiment, users call into a call center to *access VPages* using a VCode. The call center processes a request through a voice browser module that uses speech recognition and text to speech modules *to interpret user requests* for additional VPages or transactions (column 10, lines 14-36). According to another embodiment, a more *interactive approach* may be used, for example, the VPage corresponding to the Washington Monument VCode may provide the user with general information that relates to the Washington Monument. Then, the system may say "What else would you like to know about the Washington Monument?" The user could then input a voice response, which is then translated using speech recognition, natural language query technology, and/or artificial intelligence to determine the best fit of the information that is available. For example, the user may say "I would like to purchase a miniature copy of the Washington Monument" and accordingly the user would be passed to the transaction processing system to perform that task. Similarly, the user may say "I am at the Washington Monument,

Art Unit: 2142

how do I get to Arlington, Virginia?" In response to that request, the VPage may transfer the user to a VPage providing directions from downtown Washington to Arlington, Va. (see column 9, lines 24-41).

The reference Saylor in the above citation teaches an interaction dialog (e.g. the system requests or prompts, the user replies to the request or prompt and the system provides requested information) supported by a number of modules on the VNAP including a voice browser. Thus, Saylor discloses a computing system comprising a voice (conversational) browser to provide a user interface that adapts an interaction dialog between the conversation portal system and a client according to one modality, as claimed.

Appellant's arguments that the voice browser in the Saylor reference does not provide the functionalities of a quote (p. 10 or brief) "a conversational (multi-modal) browser" have been considered. However, this language is not claimed, namely, a conversational (multi-modal) browser. For the reason and the portion cited above, the reference Saylor teaches *a computing system comprising a conversational browser to provide a user interface that adapts an interaction dialog between the conversation portal and a client according to one modality*. Further, in the Cohen reference, browser (10) presents a page rendered in two modalities (e.g. graphically displayed hypertext links and voice links or both graphical and audio data and both an audio indication and a hypertext link will be available).

**A.2 Regarding claims 3 and 25, rejected as being unpatentable over Saylor in view of Cohen in further view of Cobbley, it is argued (p. 12 of brief) that the applied references do not teach:**

(i) a multi-format is a modality-independent format, because according to Appellant, improper claim interpretation (or utter lack of interpretation) as to what it meant by "modality independent format" within the meaning of the claims.

For example, (according to Appellant) as explained on pages 15-17, multi-modal content can be implemented in a modality-independent, single authoring format which characterizes the dialog interaction with the user and provide abstract representation of the dialog independently of the characteristics and UI offered by the device or application rendering the presentation material. A multi-modal document authored in a modality independent format can be processed by a conversational



Art Unit: 2142

(multi-modal) browser to provide tight synchronization between the different views supported by the multi-modal browser, where the modality-independent format can be automatically transcoded to the modality or modalities supported by the particular client browser or access device. Appellant distinguishes a multi-modal document having a modality independent format from a multi-modal document that is implemented by embedding in a single document, markup elements from each of a plurality of represented/supported modalities (e.g. VoiceXML and HTML tags), and using synchronizing tags to synchronize the different ML content (i.e., to synchronize an action of a given command in one modality with corresponding actions in the other supported modalities) on an element-by-element basis (See, e.g., page 17, lines 15 - p. 18, line 2). In view of the above, the Examiners reliance on Savior (Col. 14, lines 46-60) is misplaced. Saylor simply discloses that voice-based content can be formatted according to TML, voiceXML, WML, but clearly these formats are not modality-independent formats, but rather are modality-specific formats.

In response to the above-mentioned arguments, Appellant's interpretation of the invention and the applied reference(s) has been fully considered.

Pages 17, lines 15 - p. 18, line 2 of the specification have been reviewed. In this case the invention discloses (**verbatim**):

In another embodiment, a multi-modal CML document may be implemented by incorporating a plurality of visual and aural markup languages (i.e. a CML document that comprises sub-documents from different interaction modalities). For example, a CML document may be implemented embedding in a single document, markup language elements from each of a plurality of represented/supported modalities (e.g. voiceXML and HTML tags), and using synchronizing tags to synchronize the different ML content (i.e. to synchronize an action of a given command in one modality with corresponding actions in the other supported modalities) on an element-by-element basis. These techniques are described, for example, in the above-incorporated application International Appl. No. PCT/US99/23008, as well as US Serial Number 09/507,526 filed on February 18, 2000, entitled: System and Methods For Synchronizing Multi-Modal Interactions," which is commonly assigned and fully incorporated herein by reference.

The Boards attention is directed to the fact that there is no recitation of the argued claimed term "modality-independent format". The written disclosure cited by Appellant has not set forth "with

Art Unit: 2142

reasonable clarity, deliberateness, and precision” a controlling definition to the claimed term “modality-independent format” other than what is described on the above-mentioned portion, namely, that an multi-modal CML document *may be implemented* where *for example*, a CML document *may be implemented* embedding in a single document, markup language elements from each of a plurality of represented/supported modalities (e.g. voiceXML and HTML tags).

Thus, because there is no explicit, deliberate and precise definition for the claimed term “modality-independent format” this term will be given the broadest reasonable interpretation (i.e. a visual and voice format) and will not be limited to other embodiment(s) in the specification (see MPEP 2111.01), as required.

Hence, argued claim 3 limitation, wherein the multi-modal format is a modality-independent format, will be interpreted as, visual and voice format.

Saylor discloses that the pages rendered or presented to the user are stored in an XML-based voice file format (e.g. TML or VoiceXML) (see column 4, lines 51-53). Further, the content of a VPage is preferably an XML-based voice content file (e.g., TML or VoiceXML) that may be interpreted by a XML-based voice content interpreter and played for the user via the user's phone. The VNAP may store XML-based voice content VPages, each having a corresponding VCode and may also store an index of other information (in other formats), where selected index entries may be provided to the user (e.g., via a voice menu). If the user is interested in one of the indexed entries, the VNAP retrieves the information corresponding to the selected indexed entry and provide it to the user via the user's phone. For example, the VNAP may pass the information through a text-to-speech engine to create a sound file and play the sound file for the user via the user's phone. Alternatively, or in addition thereto, text or other content may be displayed on or output to the user's phone or other terminal device. Dynamic VPage generation may also be provided where VPage content is generated on the fly from other content (see column 8, lines 14-36).

Cohen teaches as another approach to providing access from the browser controller 102 (in the browser 10) to a Web page or automatic speech recognition (ASR) system (for example), one can include an interface using the so-called "IP telephony" protocols. As is well known, IP telephony allows simultaneous transmission of both voice and digital data. Alternatively, a parallel telephone line and Internet connection can be provided to emulate IP telephony. Yet another alternative allows the use of extensible Markup Language (XML) or another similar

Art Unit: 2142

voice/data protocol, such as Voice XML (VoxML or VXML) or HTML extensions, to provide Internet access to a PSTN application such as the browser controller 102 (column 9, lines 31-42).

Thus the applied prior reference teaches the claimed “**modality-independent format**”, namely, a page that *may be implemented using* markup language elements from each of a plurality of represented/supported modalities (e.g. voiceXML and HTML tags).

Appellant’s arguments that improper claim interpretation (or utter lack of interpretation) as to what it meant by “modality independent format” within the meaning of the claims has been applied and that the cited specification defines the claimed term and should be interpreted as characterized by Appellant has been considered but in this instance there seems to be no support for such characterization and that only a definition set forth with reasonable clarity, deliberateness, and precision will control the definition to the claimed term, as mandated.

**A.3 Regarding claim 34, rejected as being unpatentable over Saylor in view of Cohen in further view of Cobbley, it is argued (p. 14 of brief) that the applied references do not teach:**

**“wherein the at least one format comprises a multi-modal format that can be rendered in two or more synchronized modalities”,** according to Appellant for the reasons given for claims 1 and 23.

In response to the above-mentioned argument, at this outset, because the Appellant relies on the reason presented on claims 1 and 23, as the basis for arguments with respect to claim 34. Based on this rationale, examiner’s response to arguments and/or rebuttal presented in response to arguments regarding claims 1 and 23, are hereby incorporated by reference.

**B.1 Regarding claims 1, 23 and 34, rejected as being indefinite under 35 USC 112, second paragraph for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.**

In response to the above-mentioned argument, Appellant’s clarification (p. 15-16 of brief) that claimed term multimedia does include audio has been considered. This clarification

Art Unit: 2142

on the record obviated previously raised 112 rejection which is hereby withdrawn. Thus, no longer an issue to be presented for the Board's consideration.

**(11) Related Proceeding(s) Appendix**

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

Art Unit: 2142

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

*Beatriz Prieto*  
BEATRIZ PRIETO  
PRIMARY EXAMINER

bp  
December 7, 2006

**Conferees:**

*THONG VU*  
~~Abdullahi E. Salad (Prim. Ex.)~~ *Thong*

Andrew T. Caldwell (SPE)

THONG VU  
PRIMARY EXAMINER  
TECHNOLOGY CENTER 2100

*Andrew Caldwell*  
ANDREW CALDWELL  
SUPERVISOR PATENT EXAMINER

**Correspondence Address of Record:**

F. Chau & Associates, LLC.  
130 Woodbury Road  
Woodbury, New York 11797